International Journal of Applied Earth Observation and Geoinformation 23 (2013) 81–94

Contents lists available at SciVerse ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag

Mapping the distribution of the main host for plague in a complex landscape in Kazakhstan: An object-based approach using SPOT-5 XS, Landsat 7 ETM+, SRTM and multiple Random Forests

L.I. Wilschut[a,b,*], E.A. Addink[a], J.A.P. Heesterbeek[b], V.M. Dubyanskiy[c,d], S.A. Davis[e], A. Laudisoit[f,g], M. Begon[f], L.A. Burdelov[d], B.B. Atshabar[d], S.M. de Jong[a]

[a] Utrecht University, Department of Physical Geography, Heidelberglaan 2, PO Box 80115, 3508 TC Utrecht, The Netherlands
[b] Utrecht University, Faculty of Veterinary Medicine, Yalelaan 7, 3584 CL Utrecht, The Netherlands
[c] Stavropol Plague Control Research Institute, Sovetskaya 13-15, Stavropol 355035, Russian Federation
[d] Anti-Plague Institute, M. Aikimbayev's Kazakh Science Centre for Quarantine and Zoonotic Diseases, 14 Kapalskaya Street, Almaty 050074, Kazakhstan
[e] RMIT University, School of Mathematical and Geospatial Sciences, Melbourne, Victoria 3000, Australia
[f] University of Liverpool, Institute of Integrative Biology, Crown Street, Liverpool, UK
[g] University of Antwerp, Department of Biology, Groenenborgerlaan 171, B-2020 Antwerpen, Belgium

ARTICLE INFO

Article history:
Received 17 August 2012
Accepted 30 November 2012

Keywords:
Object-based image analysis
Stratification
Landscape epidemiology
Vector-borne disease
Zoonosis
Yersinia pestis
Great gerbil
Desert environment

ABSTRACT

Plague is a zoonotic infectious disease present in great gerbil populations in Kazakhstan. Infectious disease dynamics are influenced by the spatial distribution of the carriers (hosts) of the disease. The great gerbil, the main host in our study area, lives in burrows, which can be recognized on high resolution satellite imagery. In this study, using earth observation data at various spatial scales, we map the spatial distribution of burrows in a semi-desert landscape.

The study area consists of various landscape types. To evaluate whether identification of burrows by classification is possible in these landscape types, the study area was subdivided into eight landscape units, on the basis of Landsat 7 ETM+ derived Tasselled Cap Greenness and Brightness, and SRTM derived standard deviation in elevation.

In the field, 904 burrows were mapped. Using two segmented 2.5 m resolution SPOT-5 XS satellite scenes, reference object sets were created. Random Forests were built for both SPOT scenes and used to classify the images. Additionally, a stratified classification was carried out, by building separate Random Forests per landscape unit.

Burrows were successfully classified in all landscape units. In the 'steppe on floodplain' areas, classification worked best: producer's and user's accuracy in those areas reached 88% and 100%, respectively. In the 'floodplain' areas with a more heterogeneous vegetation cover, classification worked least well; there, accuracies were 86 and 58% respectively. Stratified classification improved the results in all landscape units where comparison was possible (four), increasing kappa coefficients by 13, 10, 9 and 1%, respectively.

In this study, an innovative stratification method using high- and medium resolution imagery was applied in order to map host distribution on a large spatial scale. The burrow maps we developed will help to detect changes in the distribution of great gerbil populations and, moreover, serve as a unique empirical data set which can be used as input for epidemiological plague models. This is an important step in understanding the dynamics of plague.

1. Introduction

Plague, a disease mainly spread by rodents, was responsible for the deaths of nearly one-third of the European human population during a pandemic in the Middle Ages (Gage and Kosoy, 2005; Haensch et al., 2010). Plague is caused by the bacterium *Yersinia pestis* and can infect over 200 mammal species ('hosts'), such as prairie dogs in the United States (Collinge et al., 2005), black rats in Madagascar (Keeling and Gilligan, 2000) and great gerbils in Kazakhstan (Davis et al., 2004). Plague is a vector-borne disease, i.e. the disease is transferred from host to host by a vector, fleas in the case of plague. In the last decades, plague, along with several other vector borne-diseases, has been resurging (Gubler, 1998). It causes

* Corresponding author at: Utrecht University, Department of Physical Geography, Heidelberglaan 2, PO Box 80115, 3508 TC Utrecht, The Netherlands.
E-mail address: l.i.wilschut@uu.nl (L.I. Wilschut).

0303-2434/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jag.2012.11.007

about 2000 diagnosed human cases annually, of which the majority occur in Africa (World Health Organization, 2005). The disease is still considered dangerous; it is one of three diseases for which notification to the WHO is obligatory (World Health Organization, 1983).

Much of our knowledge of disease dynamics in natural populations is gained from investigations using epidemiological models (Hudson et al., 2001). For example, population size or density often must surpass a certain threshold in order for a disease to invade and/or persist (Lloyd-Smith et al., 2005), and higher population density typically increases the transmission rate (Begon et al., 2002). Moreover, the spatial distribution of hosts may determine local host connectivity, which when increased enhances the likelihood of successful disease spread (Keeling, 1999). When host distributions are patchy, increasing 'patchiness' may either increase or decrease the chance that an endemic infectious disease will persist (Hagenaars et al., 2004; Keeling and Gilligan, 2000; Jesse and Heesterbeek, 2011; Park et al., 2002). Thus, model studies suggest that the spatial distribution of hosts influences disease dynamics in different ways, but empirical data on the spatial distribution of infected and susceptible hosts on a population scale is scarce.

The distribution of animals is increasingly mapped by earth observation and remote devices on several spatial scales. On small scales, GPS-receivers are being used to track the animals. On a larger scale, remote sensing is used: several examples exist of host-habitat mapping (Bogh et al., 2007; Kalluri et al., 2007; Estrada-Peña, 2003). Such remote sensing based maps provide useful information on the characteristics of a focus (i.e. an area where the disease causing agent and its associated vector and/or host are present). As a consequence of rapid technological advancements in recent years, high spatial resolution imagery now offers the opportunity to map actual animal distributions, provided that a feature associated with the animal, like a burrow, can be identified on satellite images.

Burrows are constructed and used by certain rodents for sleeping, nesting and food storage. Burrowing rodents may diminish vegetation cover on and around their burrows such that it becomes possible to see them on satellite images, as is the case for wombats (*Vombatus ursinus*) in Australia (Löffler and Margules, 1980) and great gerbils (*Rhombomys opimus*) in Central Asia (Addink et al., 2010).

Great gerbils, social rodents that are numerous in (semi-)deserts in Central Asia, are important hosts of zoonotic infectious diseases, such as tularemia, cutaneous leishmaniasis and bubonic plague (Gage and Kosoy, 2005; Yaghoobi-Ershadi and Javadian, 1996; Atshabar et al., 2010). Plague in the great gerbil has been monitored extensively in Kazakhstan since the 1940s, although budget cuts since the 1990s have led to a severe decrease in plague surveillance (Gage and Kosoy, 2005; Ouagrham-Gormley et al., 2008). From this monitoring, we know that plague is endemic in 20 foci in Kazakhstan, the majority being steppe and desert foci (Atshabar et al., 2010). In each of these foci, plague monitoring, i.e. testing of rodents and fleas for plague, has been carried out under the supervision of Anti-Plague Stations. The finest spatial scale on which monitoring is carried out is the so-called sector, an area, defined by a latitude-longitude grid, of approximately 9.3 km × 9.8 km.

The frequent plague epizootics (epidemics in animal populations) do not result in massive die-offs of great gerbils, because the great gerbils are – in general – quite resistant to the disease (Gage and Kosoy, 2005; Biggins and Kosoy, 2001). The occurrence of these epizootics is related to the abundance of both host and vector. A host abundance threshold was first discovered by Davis et al. (2004) where abundance was measured as the percentage of burrows occupied by family groups. The plague threshold model was later improved by including the abundance of fleas (Reijniers et al., 2012). In both models burrow (density) maps form an important starting point for the prediction of plague.

In a pilot study by Addink et al. (2010), great gerbil burrows were successfully identified in an area of 6 km × 10 km using a Quickbird image with a spatial resolution of 2.4 m. Although this study offered a convincing proof-of-concept, burrows were only mapped in one landscape type. As the distribution of great gerbils is related to the landscape by food availability, local climate and competition with other species, the challenge is to map the great gerbil abundance across different landscapes. Moreover, the area investigated in the pilot study was smaller than the area of smallest scale plague monitoring by the Anti-Plague stations. Therefore, the relation between plague dynamics on the one hand and host abundance and structure on the other cannot be studied using these data. Mapping the great gerbil burrows over large areas, covering several plague monitoring units, and across several landscape types, will offer the opportunity to study the relation between landscape, great gerbil distribution and plague occurrence.

This paper focuses therefore on classifying burrows in two areas, of 60 km × 60 km and 60 km × 85 km, covering landscape types of fluvial and aeolian origin. The objectives are:

(1) To identify great gerbil burrows by semi-automated classification on high resolution SPOT-5 XS images across different landscape types and evaluate the accuracy.
(2) To construct a map of landscape units representing the variability and spatial structure of the landscape in this plague focus.
(3) To stratify the SPOT-5 XS images on the basis of these landscape units, and subsequently classify burrows per landscape unit, using local training data.

The methodology used to map the spatial distribution of burrows was as follows: landscape units were created on the basis of Landsat 7 ETM+ and SRTM data, using object-based analysis ($4.1). Then, the SPOT images were segmented and the burrows were classified, based on unstratified and stratified Random Forests ($4.2). Finally, the accuracy for both approaches was calculated ($4.3).

2. Study area

The study area is a plague focus located in Eastern Kazakhstan, south of Lake Balkhash (Fig. 1), composed of fluvial and aeolian deposits. The area measures approximately 250 km × 200 km and encompasses a delta system developed by the Ili River. As the course of the Ili River has shifted over time from NNW to NW, several abandoned river branches can still be recognized in the landscape (Fig. 1). Large dune fields have formed along and on top of the abandoned floodplain. The irrigated agricultural area north of the town Bakanas was masked out.

Soils are sandy, with variable clay and low organic matter content: in the abandoned river beds gravel to sandy material is found, while further from the river branches, soils are more clay-rich or have dunes formed on top. In some areas soils are highly saline. The climate in the study area is strongly continental. Mean temperatures fluctuate dramatically, daily and within seasons, ranging between −40 °C in winter to +40 °C in summer (Suslov, 1961). Precipitation is less than 200 mm per year and falls primarily in winter (as snow) and spring (Suslov, 1961; Propastin, 2008). In spring snow melts quickly and forms small lakes (called takirs) in topographic depressions (Laity, 2008). Once dried up, these takirs are recognized easily because of their high albedo. Vegetation cover is variable, but shows, as well as a relation to soil moisture content, a declining trend in the direction of Lake Balkhash and with distance from the Ili River, with vegetation gradually changing from larger shrubs and reed grass thickets to lower shrubs and ephemeral grasses. Close